DOCUMENT RESUME

ED 423 280 TM 029 085

AUTHOR Dunn, Mary Anna
TITLE The Evaluation of Alternative Schools in Research and Practice.
PUB DATE 1997-03-07
NOTE 33p.; Paper presented at the Annual Meeting of the Virginia Educational Research Association (30th, March 7, 1997).
PUB TYPE Information Analyses (070) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Anthropology; Educational Innovation; Educational Practices; *Educational Research; Elementary Secondary Education; *Evaluation Methods; *Nontraditional Education; *Program Evaluation; *Qualitative Research

ABSTRACT

Although the demand for formal evaluation of alternative schools is greater than for conventional schools, the unique characteristics of these schools make evaluation especially challenging. Use of traditional instruments and methods may yield misleading results. Evaluators must understand the issues associated with the evaluation of alternative schools in order to design and conduct appropriate evaluations. This paper reviews literature related to the evaluation of alternative schools and innovative programs in order to examine the special problems associated with evaluating these schools. Examples of foci and methods are included along with recommendations for successful evaluations. Today program evaluation has emerged as a sophisticated field in its own right. The qualitative methods of anthropology have been introduced into educational research and evaluation, and the tools are available to conduct evaluations suited to the unique characteristics of alternative schools. (Contains 3 tables and 34 references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 423 280

The Evaluation of Alternative Schools in Research and Practice

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Mary Anna Dunn

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Mary Anna Dunn
University of Virginia

A Conference Paper for the 30th Annual Meeting
Virginia Educational Research Association
March 7, 1997

TM029085

The Evaluation of Alternative Schools in Research and Practice

Mary Anna Dunn

Abstract

Although the demand for formal evaluation of alternative schools is greater than for conventional schools, the unique characteristics of these schools make evaluation especially challenging. Use of traditional instruments and methods may yield misleading results. Evaluators must understand the issues associated with the evaluation of alternative schools in order to design and conduct appropriate evaluations. This paper reviews literature related to the evaluation of alternative schools and innovative programs in order to examine the special problems associated with evaluating alternative schools. Examples of foci and methods are included along with recommendations for successful evaluations.

Introduction

In 1947, the New York Supreme Court upheld Teachers College's decision to close their experimental, unit-centered Lincoln School. The school had failed to provide objective evidence that it had "functioned to the fullest extent for experimentation" (Teachers College Record in Tanner and Tanner, 1990, p. 171). While Lincoln was able to use standardized test scores to demonstrate acceptable levels of student achievement, no one conducted studies that would "measure the additional kinds of learnings and benefits purportedly afforded students by the new curriculum" (Tanner and Tanner, p. 171.) Tanner and Tanner believe that the school's failure to evaluate their innovative program contributed to its demise.

Today the evaluation of alternative schools is still problematic. While the innovative characteristics of alternative schools intensify the need for evaluation, they may also render evaluation design especially challenging. The most unique characteristics of alternative schools are not easily detected using traditional methods and instruments. Evaluators must understand the issues associated with the evaluation of alternative schools in order to design evaluations sensitive enough to detect and assess the phenomena of these schools.

Definition of the term "alternative school" is complex and controversial (Ascher, 1991; Barr, 1981; Chenoweth, 1984; Raywid, 1991; Sweeney, 1991). In the literature reviewed for this paper, usage ranges from describing court-referred disciplinary programs

to describing overwhelmingly popular public schools of choice. Terminology such as "basic skills", "enrichment", "open school", and "individualized instruction" are all used to describe curriculum and instruction. Some schools are housed separately, some are "schools-within-a-school", some are city-wide "schools-without-walls". But all these schools and programs have in common the intended implementation of a program that can be described as "distinctively different from conventional programs" (Phi Delta Kappa, 1991, p. 295).

Because of the scarcity of literature directly tied to the evaluation of alternative schools, I have included in this review literature on the evaluation of innovations. In both instances, it is the evaluator's task to describe and assess something that is distinct from what is typically offered in conventional classrooms. Lessons learned in the evaluation of innovative programs may be applied to the evaluation of alternative schools.

The Need for Evaluation

Evaluations of conventional schools are generally limited to comparisons of standardized test scores. Formal evaluations of conventional schools are rarely undertaken; formal evaluations of alternative schools may be much more frequently undertaken (Kleinbard, 1983; Raywid, 1983). In her massive survey of public schools of choice, Raywid found that 85% of alternative schools surveyed reported undergoing regular formal evaluation.

"Alternative schools are constantly being tested and judged," (Raywid, 1983, p. 686) and may have to prove their right to exist. Often evaluations result from external

pressures for alternative schools to demonstrate their worth to school boards and other funding agencies. Such a purpose is not shared with conventional schools. Although specific programs within a conventional school may be subject to termination, the right of conventional schools to exist is taken for granted by the public (Hickey, 1972; Raywid, 1983).

Although the added pressures for alternative schools to undergo evaluation may seem unfair or unwarranted, they may be advantageous to the schools (Hickey, 1972; Raywid, 1983). Raywid argues that the need for evaluation is associated with the autonomy of these schools. Looser district control leads to an increased need for evaluation.

One of the most significant potential benefits of evaluating alternative schools is continued program improvement. Formative evaluations can examine how well a school has implemented its programs, describe the effects of these programs, and recommend improvements. "First, and of perhaps of highest priority, is the purpose of internal self-improvement for the program, which in turn relates to the ongoing planning process..." (Hickey, p. 2).

An even more fundamental concern is whether the school does indeed provide a distinct alternative to what is offered by conventional schools. The evidence is that many schools referred to as alternatives do not offer distinctive programs, especially in terms of curriculum and instruction (Chenoweth, 1984; Kleinbard, 1983; Young, 1990; Cuban 1993). Chenoweth finds that labeling and other symbolic gestures in San Francisco have been substituted for implementing distinctive programs in the alternative schools in his

study. After his tenure, the first director of the San Francisco Alternative Schools Office admitted he did not know, "if anything was done differently in the alternative schools" (p. 241). Alternative programs' claims need to be validated by evaluation.

The validation of claims of innovation is important not only to insure that alternative schools do in reality offer "distinctively different" programs, but if so to support the dissemination of these innovations. One of the most important roles alternative schools can play is that of developing, evaluating, and disseminating new educational programs (Barr, 1981; Brown, 1992; Wheelock and Sweeney, 1991). "...[A]lternative Schools have been used as experimental laboratories for field-testing and validating new educational concepts ...." Barr writes, "Not since the Eight-Year Study has so much experimentation, development, and documentation occurred in public education" (p.27.) Serious and rigorous evaluation of alternative programs can function to authenticate programs and policies developed in alternative schools in the eyes of the educational mainstream.

In summary, the reasons for evaluation of alternative schools are manifold. Justification of the school's right to exist is the most unfortunate and least productive purpose. Formative evaluations of alternative schools should be conducted to understand and improve alternative schools. In addition, evaluation can serve to validate innovations for dissemination. While some schools may feel defensive about the pressure to evaluate their programs, I agree with Hickey and Raywid that these pressures can work to their advantage. If conventional schools were formally evaluated as regularly as the alternative schools in Raywid's study, they would undoubtedly be much stronger institutions.

Foci

Like all schools, alternative schools are interested in their students' achievement of basic skills, but they are at least as interested in questions asked less frequently in conventional schools. The importance of evaluating more than success in promoting basic skills acquisition is a theme that is repeatedly emphasized the literature (Brown, 1992; Chenoweth, 1983; Coppedge and Smith, 1974; Skager, 1973; Uslick and Walker, 1994).

The evaluation of alternative schools must take into account their unique purposes and interests. Hickey points out that it is a common mistake "to base evaluation of alternative schools on the basis of what traditional schools set out to do. Since most alternative schools have set out to fill a need not being met by traditional schools, evaluations must be designed on the basis of what the alternative school was designed to do" (p.2).

If the evaluation of alternative schools must be concerned with more than acquisition of basic skills, what else should it address? There is a broad variety of possible objects of evaluation (Worthen and Sanders, 1987) and the focus will be determined by the needs and interests of the schools involved. It would be beyond the scope of this paper to identify every relevant evaluation concern, however it is possible to identify broad themes into which specific concerns may be classed.

Student Outcomes

Although many authors recommend evaluations move beyond assessment of outcomes and focus at least as much on process, measurements of student outcomes still predominate. Researchers, evaluators, and school staff are alike in being particularly

concerned with assessing outcomes pertaining to affective and higher order cognitive skills including: attitudes, self-esteem, creativity, critical thinking, and organizational skills. Other questions concerning student outcomes relate to academic performance, disciplinary problems, and interpersonal relationships (Coppedge and Smith, 1975; Doob, 1977; Hickey, 1972; Skager, 1973; Weber, 1976). The Center for New Schools stresses the importance of examining student outcomes in terms of student subgroups, i.e.: "Black School-Oriented", "White Youth-Culture" (The Center for New Schools, 1972).

Teacher Attitudes and Behaviors

In a document published by the Department of Education's Program Effectiveness Panel, which reviews curricula claims for the National Diffusion Network, evidence of changes in teachers' attitudes and behaviors is regarded as acceptable evidence of an innovative program's success (Ralph and Dwyer, 1988). An evaluation that examines teacher attitudes and behaviors may be particularly useful in demonstrating whether or not something different really is occurring in alternative school classrooms.

Program Characteristics and Processes

Calling a program alternative does not make it so. One of the most important roles of evaluation may be to examine whether and to what degree an alternative school has met its goals. Chenoweth raises the concern of evaluating program distinctiveness, while Skager and educators cited in Coppedge and Smith ask, "Is the program doing what it was intended to do?" In evaluating the Chicago Metro program, the Center for New Schools goes further, asking whether the program is consistent in both its formal and hidden curricula.

Weber and Skager raise efficiency and effectiveness as evaluation concerns, although Skager holds efficiency to be a subordinate value to "goodness and rightness". It cannot be simply cost effective for children to participate in an individualized learning program, he argues. It must be demonstrated that this is a "good" way for children to learn. Skager recommends that relevant educational models and theories provide a basis for "deciding what to look for as well as how to assign values to what we observe" (p. 113).

By focusing on alignment of the institution with its alternative goals and philosophy, on program effectiveness and efficiency, and on the "goodness and rightness" of the program, evaluators include processes as well as outcomes in their scope. While this is arguably important to all program evaluations, process evaluation is particularly pertinent to the evaluation of alternative schools which claim to approach the whole process of teaching and learning in a different way.

> If we are to go on ignoring process while seeking information on only a limited set of standardized outcomes, selected on the basis of an unexamined and inarticulate core of values, the movements towards alternatives are doomed to have little support and impact. (Skager, p. 118)

Evaluation of processes is relevant not just to the evaluation of the program however, but to student evaluation as well. Because alternative schools usually emphasize the process of learning over rote memorization, Hickey and Coppedge and Smith point to the need to include process evaluation when assessing a program's effectiveness in terms of student performance. Hickey cautions that emphasizing process should not, however,

result in a de-emphasis of products. "... no program can be adequately evaluated solely on the basis of either product or process criteria" (p. 4).

Tables 1 and 2 depict specific foci of evaluations of alternative schools. Table 1 lists foci identified by Hickey in 1973. Table 2 is a summary of foci and sources accepted as evidence of success by schools for disruptive youth in Pennsylvania (Pennsylvania Department of Education, 1981). Both tables reflect alternative schools' interest in affective as well as cognitive domains, behavior, and attendance and participation. Both mention follow-up surveys as a tool or focus of evaluation.

Table 1

Foci of Evaluations Identified by Hickey

Foci of Evaluations Identified by Hickey
Community attitude towards the program
Staff attitude and attitude change over the course of the program
Parental attitude
Student attitude
Community participation
Academic achievement
Academic participation
Attendance data
Discipline and suspension figures
The extent and nature of feedback to the community
Follow-up surveys of program graduates
Holding power of the program
Changes in student family relationships
Program development and growth
Student activities outside the school
"At the elementary level, the number of kids who dash out of the room at recess"

Table 2

A Summary of Evidence of Success Reported by the Pennsylvania Department of Education

36 Schools Reporting N = number of schools reporting use of each category and information source

	STU	PAR	STA	SUR	REC	FOL	PRE	TES	CON	OTH	NG	NR
ACHIEVEMENT	0	0	0	0	1	0	1	0	0	0	2	0
ATTENDANCE	0	0	0	0	2	1	0	0	1	0	3	0
ATTITUDES	2	1	1	0	0	1	0	0	0	2	1	0
BASIC SKILLS	0	0	0	0	0	0	1	1	0	0	0	0
BEHAVIOR	1	1	1	0	1	1	0	0	0	0	2	0
COST	0	0	0	0	0	0	0	0	0	0	1	0
DROP-OUT REDUCTION	0	0	0	0	0	2	0	0	0	0	1	0
FOCUS NOT CLEAR	0	1	3	1	0	1	0	0	2	1	1	0
LOW RECIDIVISM	0	0	0	0	0	0	0	0	0	1	1	0
NONE REPORTED	0	0	0	0	0	0	0	0	0	0	0	5
RE-FUNDING	0	0	0	0	0	0	0	0	0	0	1	0
RETURN TO REGULAR PROGRAM	0	0	0	0	0	0	0	0	0	0	2	0
SELF-ESTEEM	0	0	0	0	0	0	1	0	0	0	0	0
SUCCESS AFTER GRADUATION	0	0	0	0	0	3	0	0	0	0	3	0
SUCCESS IN REGULAR SCHOOL	0	0	0	0	0	4	0	0	1	0	1	0
WAITING LIST	0	0	0	0	0	0	0	0	0	0	1	0
TOTAL	3	3	5	1	4	13	3	1	4	4	20	5

Sources of Information: STUdent **PAR**ent **STA**ff **SUR**vey **REC**ords **FOL**low-up study **PRE**test/posttest Standardized **TES**ts **CON**sultants **OTH**er No Record of Evaluation Source Not Given

Methodology

Unique Characteristics and Goals of Innovations

Dilemmas associated with using traditional tools to evaluate non-traditional schools have encumbered evaluation of alternative schools, especially before the advancement of qualitative methodology in program evaluation. Chenoweth credits the "inherent difficulty in assessing an alternative school's philosophy and goals or its distinctive instructional program," with the San Francisco school system's failure to look closely at their alternative schools (Chenoweth, 1984). When planning an evaluation of an alternative high school, Weber found it difficult to design an evaluation "which reflects the effectiveness of a new philosophy of learning or teaching" (Weber, 1976). Indeed, in spite of her best efforts to adapt a Tylerian model to the goals of an innovative program, the problems of assessing an innovative alternative program dogged her throughout the study, yielding a dearth of significant results:

> The purpose of this paper was to present a traditional-derived evaluation model for secondary education, the results of application of that model, and the implications of the application for a "better" evaluation model. The first two parts of this stated purpose have been achieved. The third and critical part is only partly addressed. (p 43).

Quantitative/Qualitative

The crux of Weber's disappointment may derive from her over-reliance on quantitative, quasi-experimental methodology:

The Evaluation of Alternative Schools

> There is little question in the research community today that the "experiment" provides the optimally sound bases for educational policy decisions. However, the feasibility of randomization, the sine qua non of experimental design, may be limited for local evaluation. In such instances the problem for the evaluator becomes one of making the best of a bad situation. Rather than testing hypotheses or estimating effects, the evaluator may have to fall back on traditional evaluation methodology ... (p. 5).

For some educational programs, limiting evaluation to experimental and quasi-experimental designs provides too narrow a view of what can be considered evidence (Ralph and Dwyer, 1988; Skager, 1973).

Of particular concern is the validity of using standardized test scores as an indicator of an alternative school's success. Standardized test scores have very important information to offer about the individual student's achievement as well as how schools are doing in comparison to other schools in the area. But reliance on standardized test scores is not likely to yield enough data about an alternative program.

While evaluations of basic skills using standardized tests may provide important details about an alternative school, they cannot paint a complete picture. Evaluations based on comparisons of standardized test scores may not reveal significant differences in spite of a program's obvious distinctiveness (Brown, 1992; Shapiro, 1973; Smith, Barr, and Burke, 1976). Even higher than average standardized test scores are not necessarily evidence of superiority, but may instead reflect attributes of the population attracted to the schools. In Chenoweth's study of alternative schools in San Francisco, standardized test

scores were traditionally higher in alternative schools than in other district schools, yet Chenoweth found little evidence that the schools were doing anything differently. Skager cites studies of the implementation of alternative schools in Sweden in which standardized test scores at alternative schools were compared to scores at conventional schools. The fact that students at the alternative schools did not score lower than those at conventional schools was taken as evidence of these schools' success, but a subsequent study revealed that the program in the alternative schools had never been implemented as intended, and the quality of instruction was poor. Standardized test scores provide information about individual student performance, not about the "quality of teaching and learning" (Stufflebeam and Webster, 1983, p. 30).

Even when used to evaluate student performance, standardized test scores are not likely to be adequate indicators of how well all of an alternative program's objectives are being met. In their evaluation of a math enhancement project, Uslick and Walker found that standardized test scores lacked validity because they were not aligned with program goals. Indeed, teachers felt a conflict between their interests in achieving the innovative program's goals and the pressure to produce acceptable scores on standardized achievement tests. Similarly, the teachers in Brown's evaluation expressed concerns that standardized tests would not immediately reflect their focus on learning processes, collaboration, and higher order thinking.

Scores on standardized tests, whether they are achievement tests used regularly by the district, or specific tests selected by evaluators to target evaluation questions, may yield some important information for comparative purposes, but they are less likely to

detect the presence and effects of distinctive programs, particularly when there are no opportunities for randomization. Reliance on quantitative data derived from standardized test scores may lead to frustration or complacency and will provide little formative information about a school.

Evaluating an innovation solely on the basis of quantitative information is comparable to placing a value on a precious stone on the basis size and weight alone. This method would fail to detect the difference between a diamond and a zircon, and similarly, purely quantitative information about innovative programs may not detect the differences between them and their conventional counterparts.

Taking the physical measurements of a precious stone is a step in the process of placing a value on it, but a jeweler must also skillfully assess the quality of a gemstone. Qualitative methods can be used to assess the essential qualities of a program that may be missed using only quantitative measures such as test scores and attendance figures.

Because schools are multi-faceted, qualitative studies may be valuable in dealing with the complexities of programs with several distinctive features that might confound an experimental design (Barr, 1981; Brown, 1992; Ralph and Dwyer, 1988; Skager, 1973). Skager objects to the experimental model of evaluation because experimentalists tend to focus on individual variables while missing the interrelationships of these variables as components of a whole. Evaluating an elementary school undergoing multiple reforms, Brown reports:

> Evaluation of the effect of the school project was very difficult because of the large number of other curricular changes that were taking place at the school at the same

time. Analysis of qualitative data provided an understanding of the effect of different components of the project from parents' and teachers' perspectives. (p. 6)

What may be most relevant to the evaluation of alternative schools is the sensitivity of qualitative methods in elucidating the essence of a program, the phenomenon that eludes quantitative studies. "Participating in the individual classroom was the only way to describe the essence of if, when, and how changes would occur", Uslick and Wheeler discovered in their evaluation of an innovative mathematics program (p. 7). Shapiro found no differences between achievement tests in an enrichment program based on open education and a comparison group, yet found striking differences in classroom observations. Conversely, Chenoweth's qualitative research on San Francisco alternative schools suggested what standardized scores could not detect: that the schools' claims of distinctiveness were more symbolic than substantial.

Qualitative research also has the advantage of yielding powerful information about the human impact of a program by allowing participants to express themselves in their own words (Patton, 1980) As part of their evaluation, Uslick and Wheeler conducted a "mini-case study" which helped them understand the frustration experienced by the teachers implementing the project. Interviews and open-ended surveys can also improve understanding of the impact alternative schools have on their students. Since alternative schools may target students who have experienced feelings of frustration in conventional programs , this information will be especially useful in understanding how these schools have achieved their goals. In addition, since the focus on individualization of many

alternative schools makes aggregation of data difficult, intensive qualitative case studies may provide a more vivid picture than quantitative data can portray.

Studies that combine quantitative and qualitative methods will provide broad perspectives and meet the varied needs of stakeholders. Concerned that combined qualitative and quantitative research might result in mediocre work Brown designed a team that included one member each for qualitative and quantitative research. The qualitative researcher collected data to analyze processes and the impact of the program on parents' and teachers' attitudes while the quantitative researcher collected data on basic skills, reading ability, and student attitudes.

Table 3, Qualitative and Quantitative Tools Used in Evaluations Reviewed

Quantitative	**Qualitative**
Aptitude Tests	Interviews
Attitude Surveys	Journals
Follow-up Questionnaires	Observations
Pre-tests and Post-tests	Open-ended Questionnaires
Standardized Achievement Tests	Video-tapes
Student Records	

Whether an evaluation is quantitative, qualitative, or mixed, one of the tools associated with qualitative methodology is especially pertinent to the evaluation of alternative schools: description. Much can be assumed about a conventional school that may not be true of an alternative program. While descriptive material is important in all evaluations, if a school is distinctive, it is particularly important that the evaluator to be clear and specific in providing descriptive information about its context, programs, goals, and outcomes (Cronbach, 1983; Shapiro, 1973; Weber, 1976) At the same time, the evaluator must avoid the trap of providing more description than assessment. In their review of 19 evaluations of alternative schools, Duke and Muzio found many evaluations so steeped in descriptive material that their "archival value thus outweighs their usefulness as input to decision making" (Duke and Muzio, 1978, p. 464).

Comparison Groups

The question of whether or not comparison groups can or should be employed in the evaluation of alternative schools is subject to debate. Two issues which may hamper the feasibility of comparison studies are randomization and identification of meaningful comparison groups.

Dysanrski argues that it *is* possible to design a randomized comparison study of alternative schools when a school has a larger pool of applicants than openings. In this case, as some students must be excluded, students may be selected randomly. Ideally, there should be twice as many applicants as openings, allowing for a 1:1 ratio of students

assigned to traditional and alternative schools. A minimum of 50% more applicants than openings is necessary, allowing a ratio of 1:2.

In Dysanrski's evaluation, assignment of students was stratified, assuring administrators of a mix of students that met their criteria. In addition, a few "wild cards" were allowed, that is to say a small number of spots were held open to allow staff the flexibility to admit of students with special needs or circumstances. (Dysanrski, 1994).

The use of lotteries, often employed to deal with the demand for spots in popular alternative schools, provides an opportunity for the types of randomized study Dysanrski advocates. This method was used to evaluate Chicago's Metro High School. Control groups were "composed of students who applied to Metro but were not enrolled after a random selection process" (Doob, 1977, p. 22). Comparisons were made in terms of achievement, attitudes towards schools and community, and success following graduation.

Weber approached the problem of randomization differently, designing a non-randomized comparison study with statistical adjustments. Students participated in the school-within-as school, known as The Community Group, voluntarily and therefore she believed students in the parent school should participate in the study voluntarily. "A random sample from the parent high school would provide no more meaningful a comparison than would be made with groups of students recruited from the parent high school" (p. 19). Weber also believed that randomly selecting students for the study would reduce post-test participation. Weber used covariance adjustments to reduce bias.

The frustrations Weber experienced in conducting this evaluation reflect the difficulties inherent in trying to construct a comparative study for an alternative school. If

the school is truly an alternative, then the evaluator risks comparing the proverbial apples and oranges.

> The problem of comparison group identification is inherent in the evaluation of "new" programs. Attempts to change the process of education and thereby redefine the goals of education make identifying a comparison group difficult. It is more difficult to identify a comparison program that agrees with the priorities shared with an innovative program. And, the more novel or innovative that the program under study is, the more difficult it becomes to find an appropriate comparison. (p. 5).

In the end, in spite of her carefully wrought study, Weber concludes, "The results of the case study here suggests that comparisons do not yield meaningful evaluation of an innovative program" (p. 41).

In addition to the difficulty of identifying comparable programs, the choice status of most alternative schools confounds the issue. "Too frequently research focuses on cognitive skills development," writes one respondent in Coppedge's and Smith's study, "And neglects recognition that students in alternative programs by natural selection are unusual. Therefore, comparative research is frequently invalid or unreliable" (p.15).

It can be argued that comparative evaluations may not be necessary or appropriate. Hamilton argues that "comparative strategies are linked most closely to the behavior of individuals" but anthropological models are more relative to the evaluation of social institutions (Hamilton, 1976, p. 84) and Kocher asserts that the evaluator should be

concerned with how a program is meeting its own goals, not how it compares to another program (Kocher, 1975).

Yet, comparisons can be very helpful in providing information about the benefits and costs of a program (Patton, 1986). Comparisons of alternative schools can and have yielded dramatic and convincing results (Doob, 1977). Without them, it would be hard to say how much difference an alternative school has really made (Duke and Muzio, 1978).

While comparisons may be neither necessary nor appropriate for all evaluations there will be times when they will make an evaluation much more relevant. Based on these readings I offer the following guidelines:

1. Comparative evaluation is only one of many tools available for evaluation, it should not be regarded as the sine qua non of evaluation, but rather selected when it is evident that it is the tool needed to investigate the questions posed.
2. Comparative evaluations should be aimed at goals that are specific to the school. For example, if an alternative school shares the goal of teaching basic skills with a comparison school, but is unique in its emphasis on improving students' attitudes towards school, then , though readily accessible, achievement test scores will yield less useful information than attitudinal measures.
3. Comparative evaluations will work best if designed for simple programs. The more complex the program, the more difficult it will be to identify an appropriate comparison group.

4. To avoid the effects of self-selection, wherever possible comparison groups should be drawn from applicants who applied, but were randomly denied admission to, the alternative school.
5. Taken alone comparative studies can be confusing or misleading. Comparative studies should be interpreted in light of other evidence about the school.

Objective-Oriented and Goal Free Evaluations

Objective-oriented evaluations, introduced by Ralph Tyler, are used to determine to what extent a school's objectives are actually being met. While some of the authors reviewed in this paper struggled with the appropriateness of the Tylerian model to the assessment of alternative schools, it should be remembered that the Tylerian model was developed to evaluate progressive schools in the highly successful Eight-Year study. I doubt it is the Tylerian model, but the application of the model that is used inappropriately in unsuccessful evaluations. Care should be take that the objectives identified for evaluation are the programs' objectives and that measurements are suited to these often unique goals.

At the same time, since the Eight-Year study, other evaluation methods have been introduced that should be considerd as well. Developed and advocated by Michael Scriven, goal-free evaluations focus on actual rather than intended outcomes. The evaluator is kept blind to the program's goals, thus the focus of the evaluation is broader and the evaluator has a greater chance of detecting unanticipated consequences as well as aspects of the program that are not objective-oriented. (Worthen and Sanders, 1987).

The Evaluation of Alternative Schools

Weber holds that it is possible to use an adapted modified Tylerian model to evaluate alternative schools but she also incorporates a goal free evaluation into her study to help understand the impact of the innovations. Hamilton argues that while evaluators may attempt to modify the Tyler/Bloom model, the goal free model is more appropriate to a curriculum with diffuse and general objectives. Duke and Muzio wonder why goal free evaluations were not included in the evaluations they reviewed since alternative schools "perhaps more than conventional schools, are likely to produce a number of unintended consequences" (p. 482).

Follow-up Studies

Among the most popular types of studies identified in both the Coppedge and Smith and in the Pennsylvania Department of Education studies were various forms of follow-up studies that examined student success on leaving the alternative program. As Duke and Muzio put it, "One of the most telling indicators of a school's effectiveness can be the success of its graduates" (p. 477). Given the limited value of standardized testing in the evaluation of alternative schools, follow-up and longitudinal studies have a promising role in assessing how well graduates are prepared.

Implementing a Successful Evaluation

While formative evaluation is crucial in the early stages of an alternative school's implementation, summative evaluations should be deferred until the school is established. Brown argues that while funding requirements often demand that evaluations of reforms be conducted in a year or two, these demands are "dysfunctional" because changes in

student knowledge may not be picked up by standardized tests so quickly. (Brown, p.4). Likewise, Kocher asserts that "new programs should be given three years to get established, during which time evaluators should provide feedback to the school as to how to improve it, not whether or not it should exist" (Kocher, p. 47).

All evaluations should be conducted by evaluators who are able to understand the nature of alternative education. In 1971, Hickey worried about "the lack of qualified evaluators who have the sensitivities and insights necessary to fully understand the concept of alternative evaluation and measure its implementation" (p. 4). A quarter of a century later, general understanding of alternative schools is still obscure enough to raise concerns. Schools will be wise to take care in choosing an evaluation team that understands alternative education, and to take responsibility themselves for educating evaluators about their programs.

Perhaps the greatest impediment to conducting a successful evaluation is the availability of sufficient time and resources. Respondents in Coppedge's and Smith's study reported that lack of sufficient time, money, and tools interfered with developing improved evaluation programs. Their research suggests that schools would conduct more and better designed evaluations if they had the resources. Although case studies are recommended for assessing innovative programs, Ralph and Dwyer acknowledge that they are "costly, complex, and time consuming" (p. 12).

Insufficient resources may lead to evaluation designs that are too narrow to detect and assess the phenomena of an alternative school. In formative evaluations, this may lead to a failure to promote more successful implementation of the program. In summative

evaluations, this may lead to the continuation of programs that have not succeeded in meeting their goals or the closing of schools that have.

Conclusions

Fifty years since the Lincoln School closed its doors, the mortality rate among alternative schools is high (Cuban, 1993). We are not locked into the evaluation methods applied fifty years ago. Today, program evaluation has emerged as sophisticated field in its own right. The qualitative methods of anthropology have been introduced into educational research and evaluation, opening a wider window upon our understanding of education programs. The tools are available to conduct evaluations suitable to the unique characteristics of alternative schools. Careful and appropriate evaluations can go a long way towards strengthening alternative schools by improving individual schools and communicating findings with other schools (Center for New Schools, 1972). If the survival rate of alternative schools is to improve, the best tools available must be employed to conduct suitable and meaningful evaluations. At the same time, research needs to be conducted into the problems and special issues that continue to be associated with the evaluation of alternative schools in order to develop the resources and methods that will improve the state of the art.

References

Ascher, C. (1992) "ERIC/CUE: Alternative schools -- some answers and questions. In Williams, S. (Ed.). The alternative school choice (pp. 5-9). Bloomington, In: Phi Delta Kappan. First Published in (1982) Urban Review, 14 65-69.

Barr, R.D. (1981) Alternatives for the eighties: a second decade of development. Phi Delta Kappan, 62 570-73.

Brown, M.J.M. (1992). Issues in educational reform: how triangulating qualitative and quantitative evaluation methods can enhance understanding. Paper presented at the annual meeting of the American Evaluation Association. Seattle, Washington.

Center for New Schools (1972). Strengthening alternative high schools. Harvard Educational Review, 42 313-350.

Chenoweth, T.G. (1984). Perspectives on the survival of alternative schools. (Doctoral Dissertation, Stanford University, 1984) Dissertation Abstracts International.

Coppedge, F.L. and Smith, G.R. (1974). Evaluation practices and preferences in alternative schools. Teacher Education Forum Series. 2, No 20. ERIC Document ED 099 329.

Cronbach, L.J. (1983). Course improvement through evaluation. In Madaus, G.F, Scriven, M. and Stufflebeam, D.L. (Eds.) Evaluation models: view points on educational

and human services evaluation. (pp. 101-115) Boston: Kluwer-Nijhoff Publishing. First published in (1963) Teachers College Record, 64 672-83.

Cuban, L. (1993) How teachers taught; Change and constancy in American classrooms, 1880-1990. 2nd Edition. New York: Teachers College Press.

Doob, H.S. (Ed.) (1977). Evaluations of alternative schools. Arlington, Va: Educational Research Service, Inc.

Duke, D. and Muzio, I. (1978). How effective are alternative schools? -- A review of recent evaluations and reports. Teachers College Review, 79, (3) 461-483).

Dysanrski, M. (1994). Implementing rigorous evaluations of education interventions: findings from two federal demonstration projects. ERIC Document Reproduction Service Number. ED 373 097.

Hamilton, D. (1976). Curriculum evaluation. London: Open Books.

Hickey, M. (1972) Evaluating alternative schools, Changing Schools: An Occassional Newsletter. ERIC Document ed 071 963.

Kleinbard, P. (1983). Alternative schools: current issues. Social Policy 13: 46-7.

Kocher, A. T. (1975). Issues in the formative evaluation of alternative schools. Paper Presented at the Annual Meeting of the American Educational Research Association, Washington, D.C. ERIC Document ED 109 173.

Patton, M.C. (1980). Qualitative evaluation and research methods. Newbury Park, Ca.: Sage Publications.

Patton, M.C. (1986). Utilization-focused evaluation. Newbury Park, Ca.: Sage Publications.

Pennsylvania Department of Education. (March 1981). Alternative education programs for disruptive youth. ED 209 770.

Ralph, J. and Dwyer, M.C., (1988). Making the case: evidence of program effectiveness in schools and classrooms. U.S. Department of Education.

Raywid, M. A. (1991). Synthesis of research on schools of choice. . In Williams, S. (Ed.), The alternative school choice (pp 133-141). Bloomington, In: Phi Delta Kappan. First published in (1984) Educational Leadership, 41, (70-78).

Raywid, M.A. (1983) Schools of choise: their current nature and prospects. Phi Delta Kappan, 64, 684-688.

Shapiro, E. (1973). Educational evaluation: rethinking the criteria of competence. School Review, . 523-549.

Sizer, T. (1991). High school reform: the need for engineering. In Williams, S. (Ed.), The alternative school choice (pp 77-81). Bloomington, In: Phi Delta Kappan. First published in (1983) Phi Delta Kappan, 40 679-683.

Skager, R. (1973). Evaluating educational alternatives. In Goodlad, J. (Ed) New directions for education: Vol 4. Alternatives in education, (pp 97-120). Dan Francisco: Josey-Bass.

Smith, V., Barr, R. and Burke, D. (1976). Alternatives in education. Bloomington Indiana: Phi Delta Kappa.

Stufflebeam, D.L. and Webster, W.J. (1983). An analysis of alternative approaches to evaluation. In Madaus, G.F., Scriven, M.S., and Stufflebeam, D.L., (Eds). Evaluation models: Viewpoints on educational and human services evaluation. (pp. 23-43) Boston:: Kluwer-Nijhoff Publishing.

Sweeney, M.E. (1991) Alternative education and "alternative" schools. In Williams, S. (Ed.), The alternative school choice (pp 11-20). Bloomington, In: Phi Delta Kappan. First published in (1988) Holistic Education Review, 1, 22-25.

Tanner, D. and Tanner, L. (1990). History of the school curriculum. New York: MacMillian Publishing Company.

Uslick, J. and Walker, C. (1994). An evaluation of an innovation: standardized test scores were not valid indicators of success. A Paper Presented at the Anual Meeting of the American Educational Research Association. New Orleans, Lousisana.

Weber, M.B. (1976). Alternative education: implications for evaluation methodology. Paper Presented at the Annual Meeting of the American Educational Reseach Association, San Fransico Ca.. ERIC Document ED 126 102.

Wheelock, A. and Sweeney, M.E. (1991) Alternative education: A vehicle for school reform. . In Williams, S. (Ed.), The alternative school choice (pp 11-20).

Bloomington, In: Phi Delta Kappan. First published in (1989) Changing Schools, 17 (4-6).

Williams, S. (Ed.), The Alternative school choice. Bloomington, In: Phi Delta Kappan (1991).

Worthen, B.R. and Sanders, J.R. (1987). Educational evaluation: alternative approaches and practical guidelines. New York, NY: Longman.

Young, T.W. (1990) Public alternative education: options and choices for today's schools. New York: Teachers College Press.